Astroinformatics 2019
Caltech, 2019 June 26

Can we predict (the future of) aperiodic sources?

Matthew J. Graham
Center for Data-Driven Discovery/ZTF, Caltech
mjg@caltech.edu

CENTER FOR DATA-DRIVEN DISCOVERY

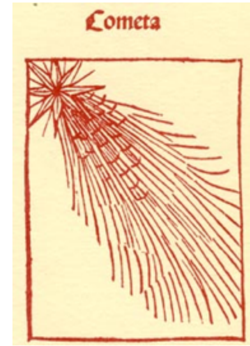
ZTF

A history of predictive modelling

164 BCE

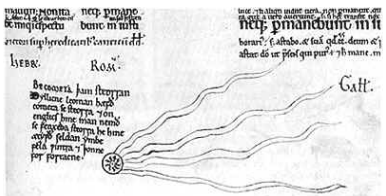
87 BCE

684



1066

1145



1301

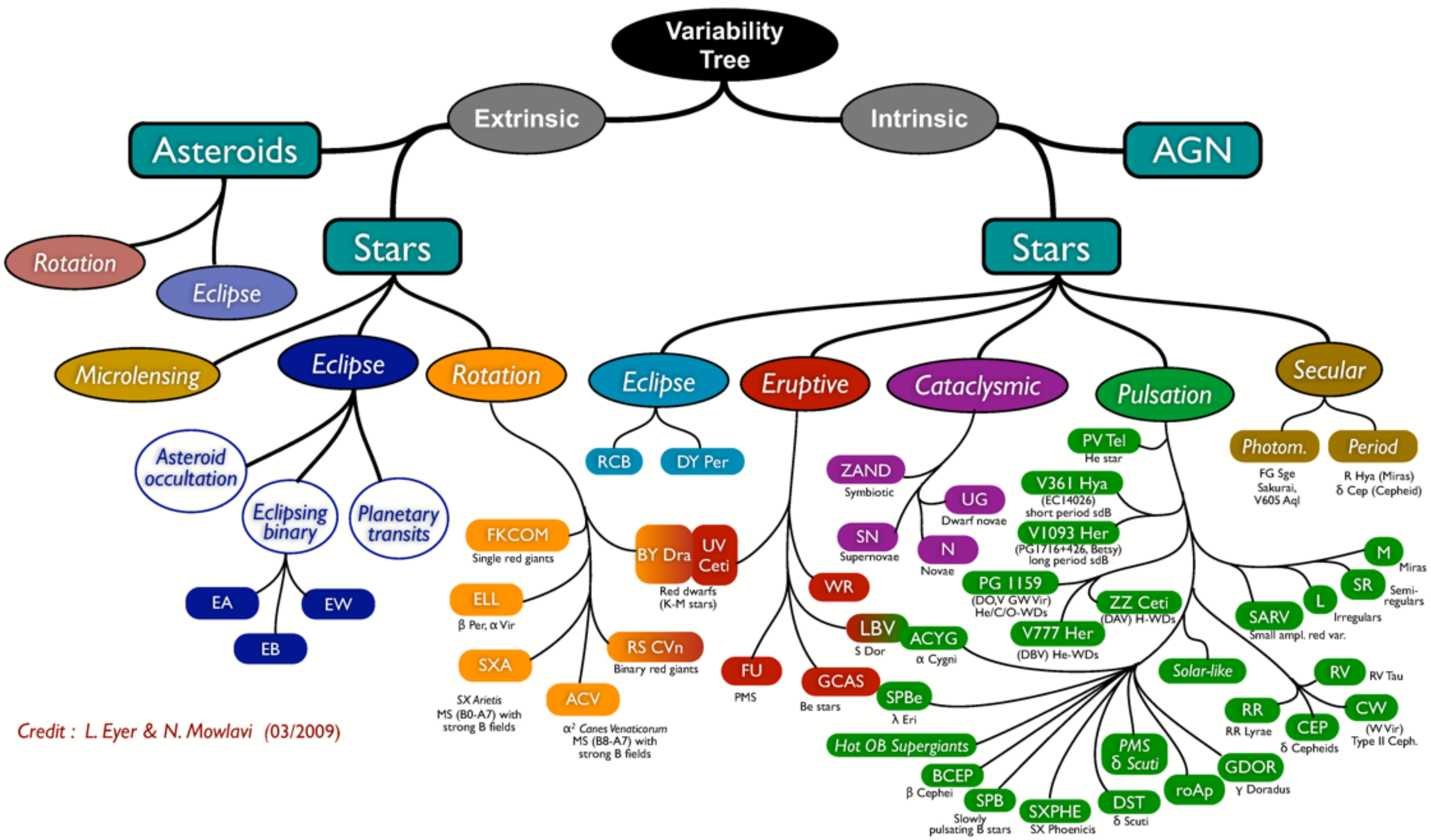
1456

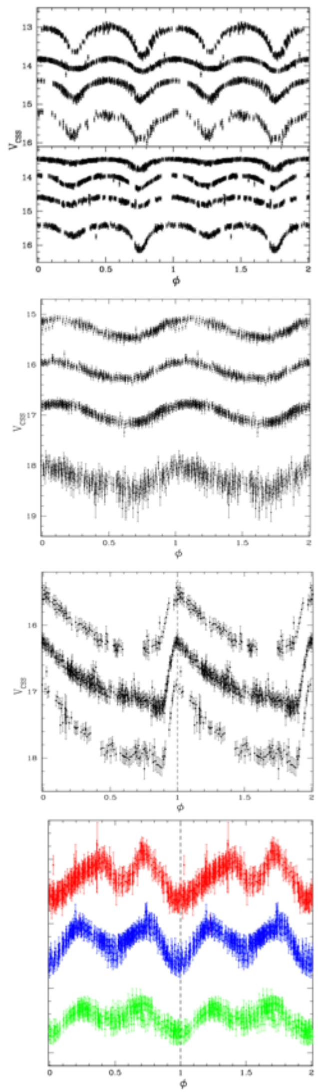
1682

1835

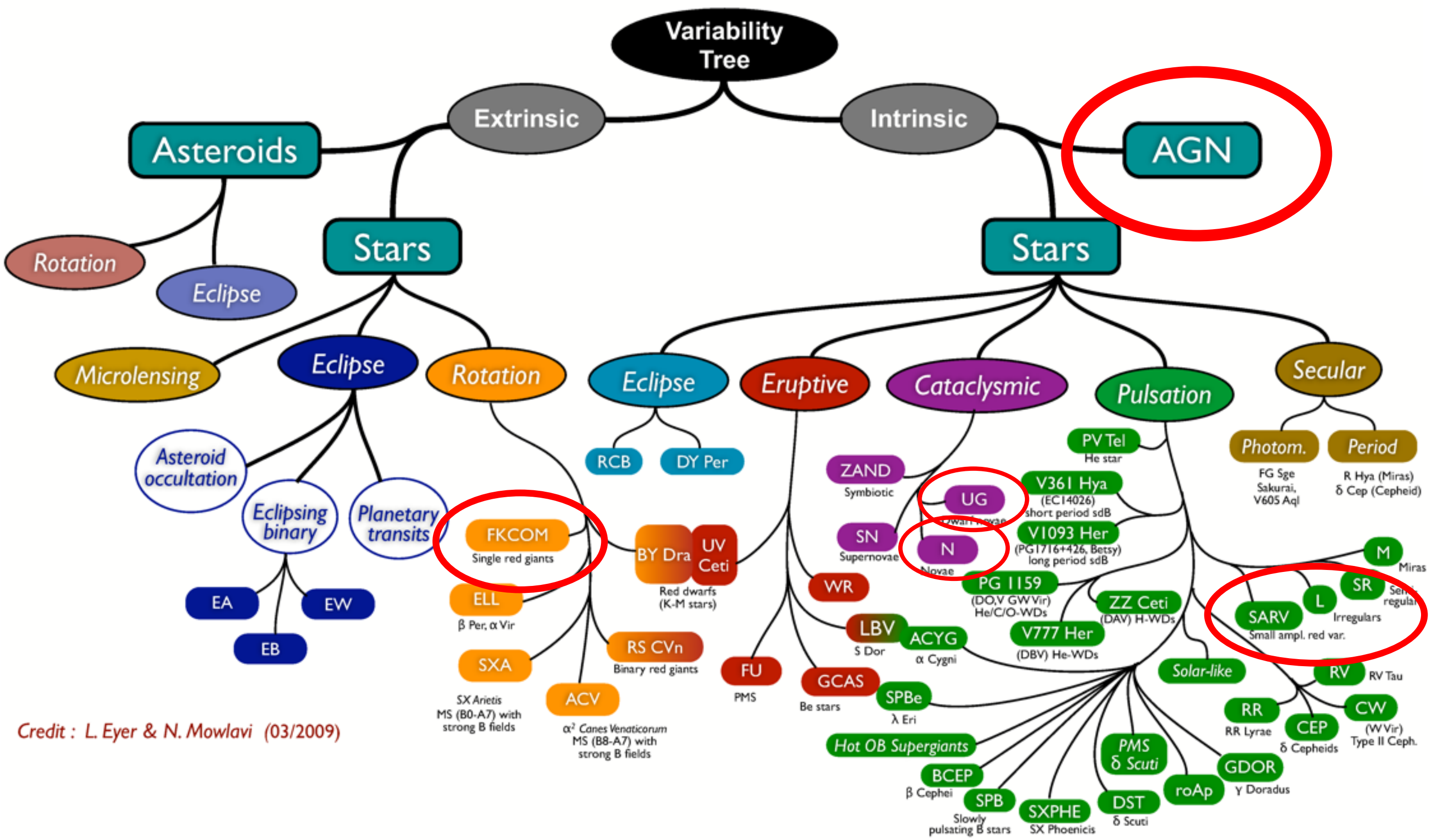
1910

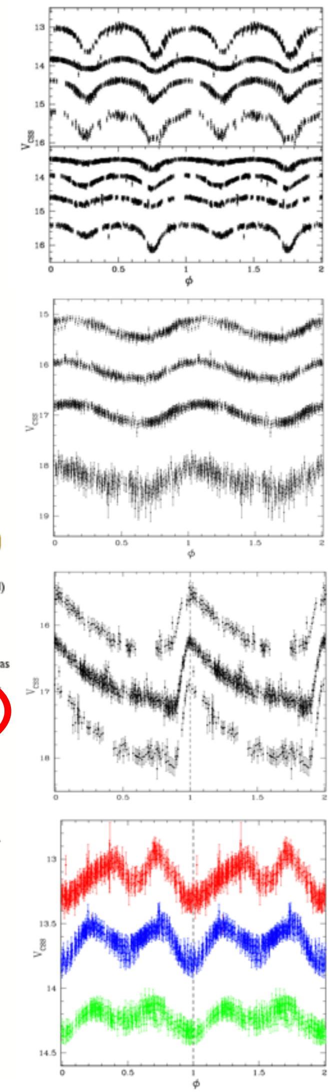
Periodic sources in astronomy





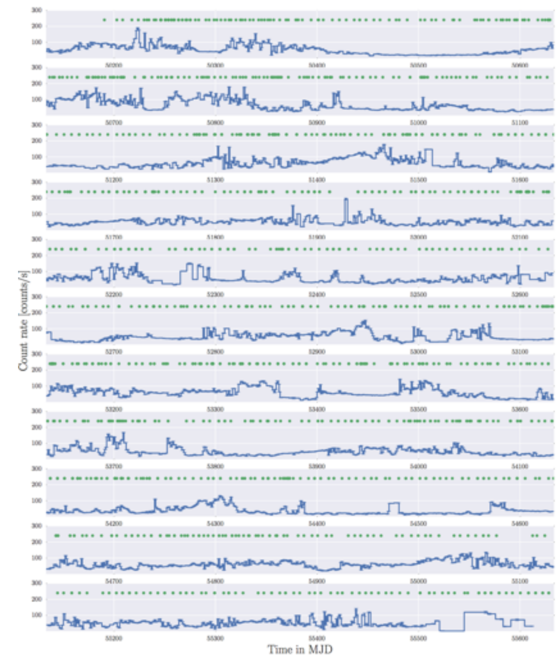
Periodic sources in astronomy





Aperiodic sources in astronomy

- Accretion systems:
  - ➢compact stellar sources to SMBHs
  - ➢YSOs
  - ➢planetary disks
  - ➢galaxy formation
- Chromospheric activity
- Boyajian's star
- Stellar associations



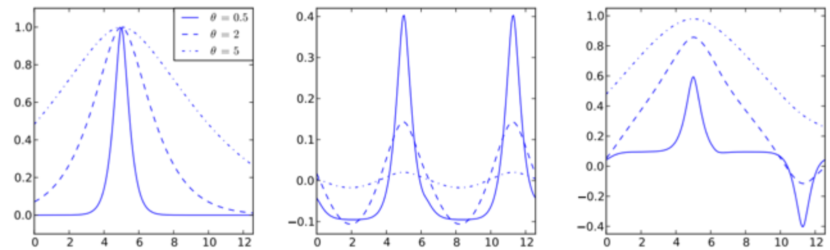
(Huppenkothen et al. 2016)

The challenge is to identify those phenomena that are unexpected for aperiodic sources (even for volatile behaviors) - periodicity

- Characterization between different processes

Measuring (a)periodicity

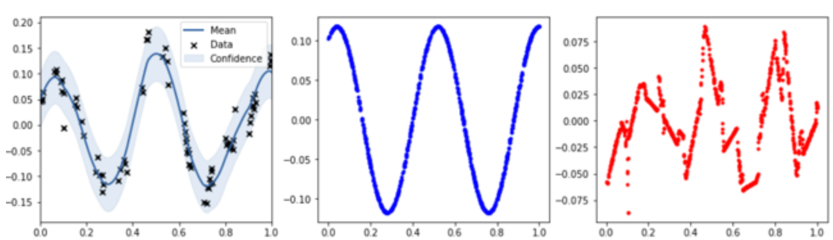
- Harmonic analysis is based on the projection of a function *f* onto a periodic basis set but problematic with finite sampling of *f*
- Durrande et al. (2016) show how a Gaussian process with a (Matérn) kernel can be projected onto sub-reproducing kernel Hilbert spaces such that: $k = k_p + k_a$

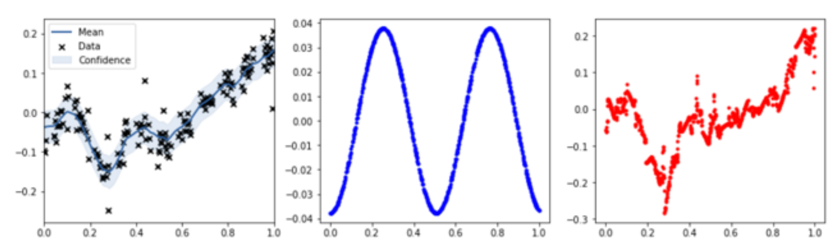


(a) Matérn kernel k. (b) periodic sub-kernel k_p. (c) aperiodic sub-kernel k_a.

- The degree of periodicity can then be estimated from sensitivity analysis:

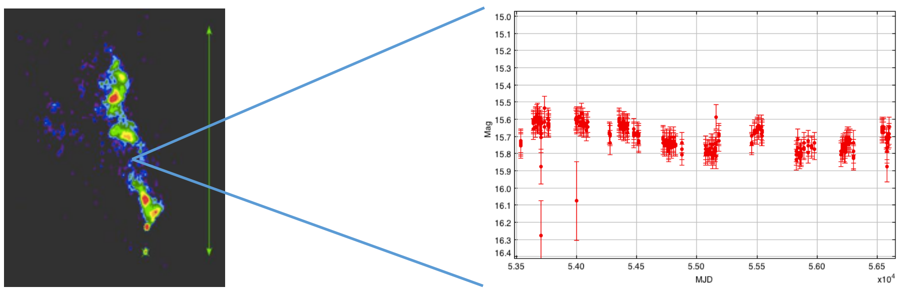
$$R = \frac{\mathrm{Var}_T\left[Z_p(T)\right]}{\mathrm{Var}_T\left[Z_p(T) + Z_a(T)\right]}$$



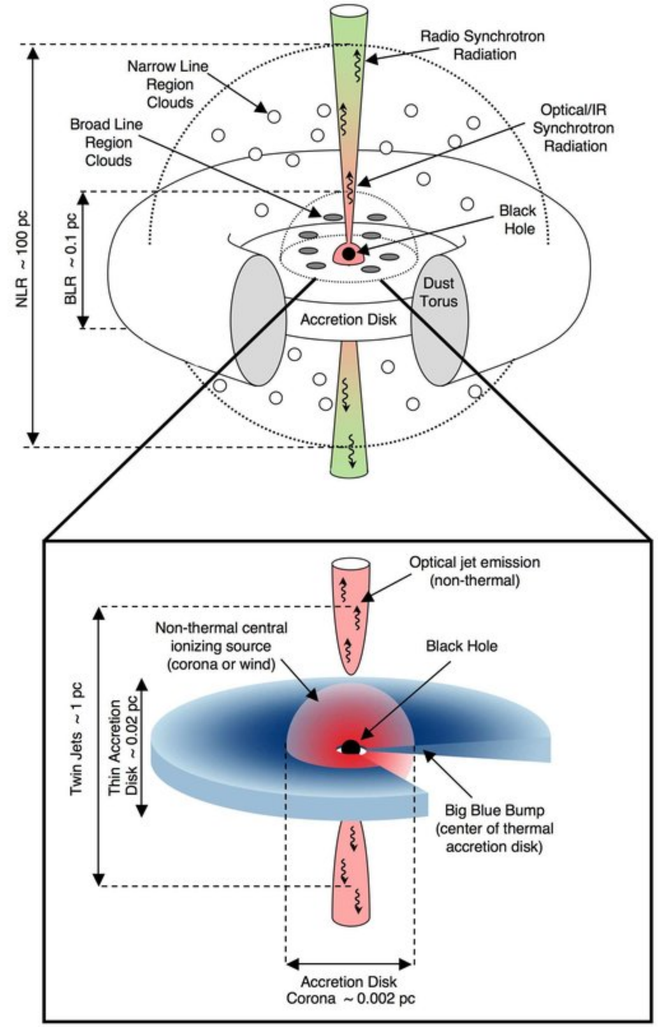


The quintessential aperiodic population

- First quasar identified 3C 48 – most striking feature was that the optical radiation varied

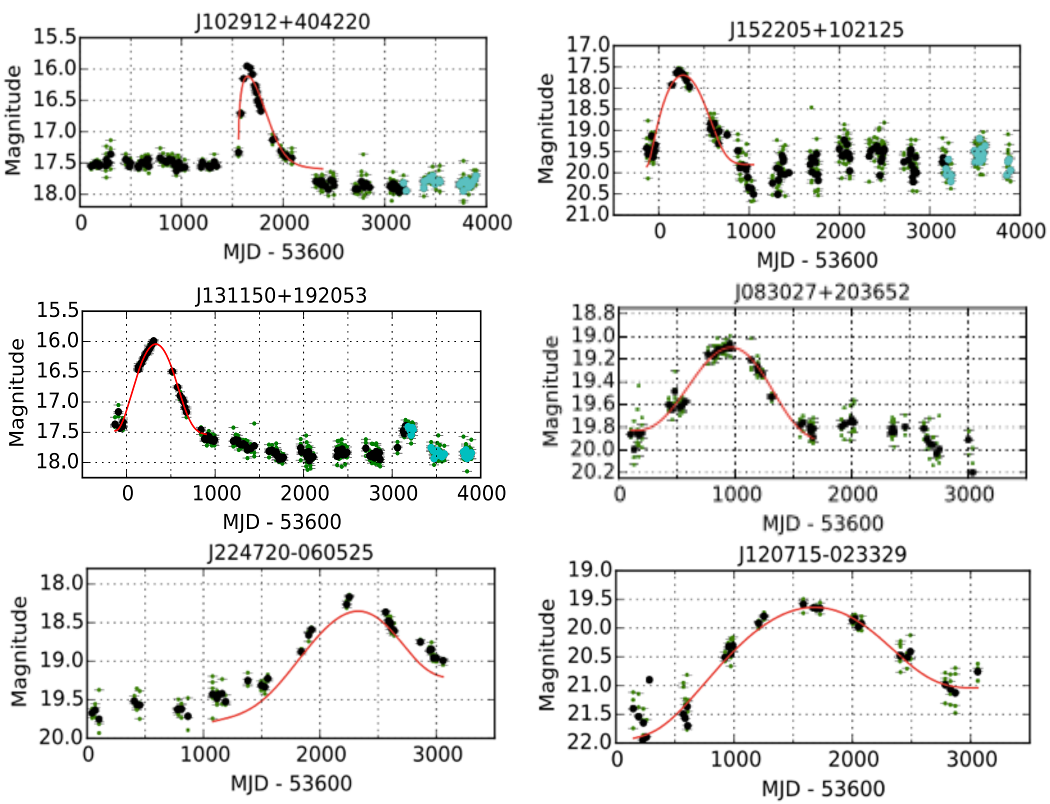


- Physical origin of photometric variability in optical/UV is unclear:
  - Instabilities in the accretion disk
  - Supernovae
  - Microlensing
  - Stellar collisions
  - Thermal fluctuations from magnetic turbulence

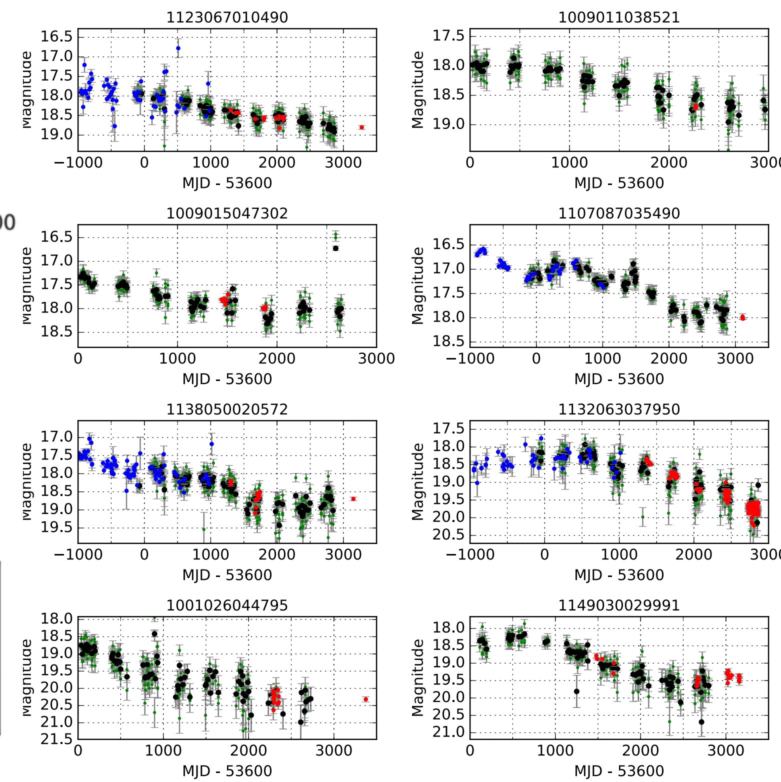


Unexpected quasar behavior

- Flaring activity
- Microlensing



- Changing-state (changing look)

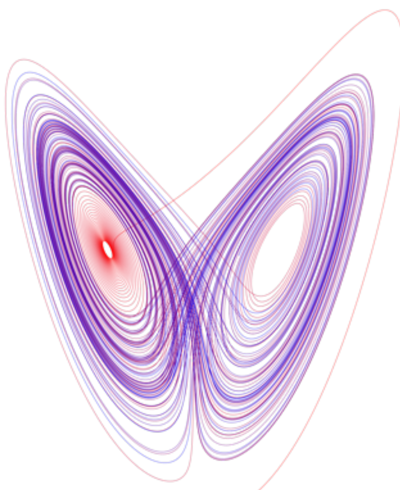


Stochastic vs deterministic

- Is it random? Or just very nonlinear? (Is it even stationary?)
- Ensemble of unknown processes means a large number of degrees of freedom => stochastic perturbation:

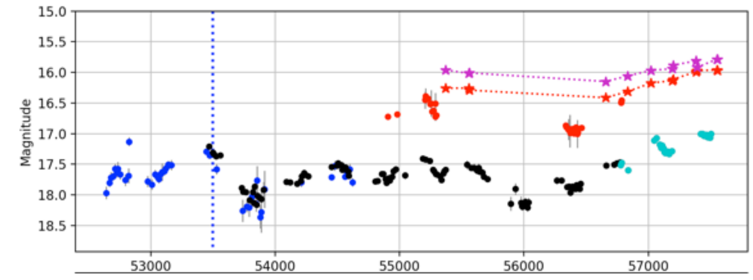
$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t$$

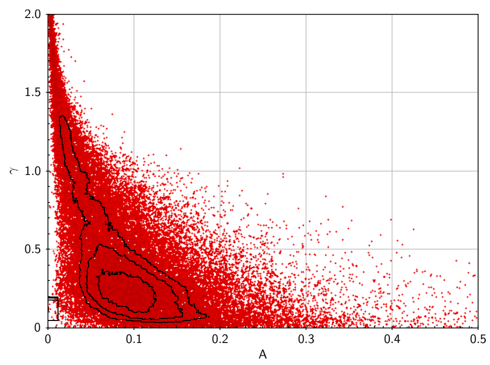
- Highly nonlinear system with bounded trajectory in phase space:

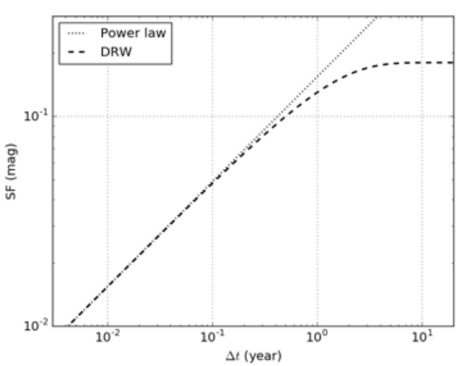


Describing quasar photometric variability

- $|\Delta m| > x$
  - DPOSS vs. SDSS (Stripe 82) vs. PS1
- Excess variability: χ^2
- Structure function
  - Variability amplitude as a function of the time lag between compared observations
  - Historic descriptor of variability and a variety of estimators
  - Not much information



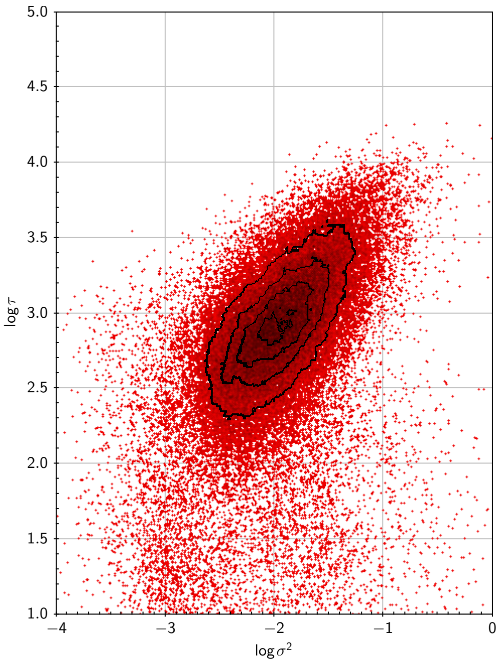


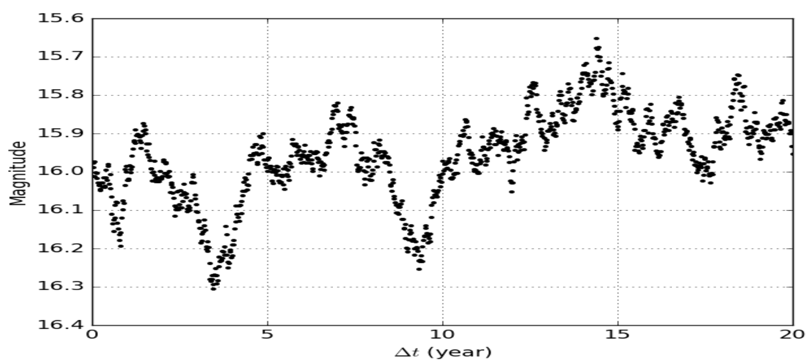


Damped random walk (DRW/OU)

$$dX(t) = -\frac{1}{\tau}X(t)dt + \sigma\sqrt{dt}\,\varepsilon(t) + b\,dt \quad \tau, \sigma, t > 0$$

- Characterized by variability amplitude and timescale
- Basis for stochastic models of variability
- Deviations noted (e.g., Mushotzky 2011, Zu et al. 2013, Graham et al. 2014)
- Degenerate model – can be best fit for a non-DRW process (Kozlowski 2016)



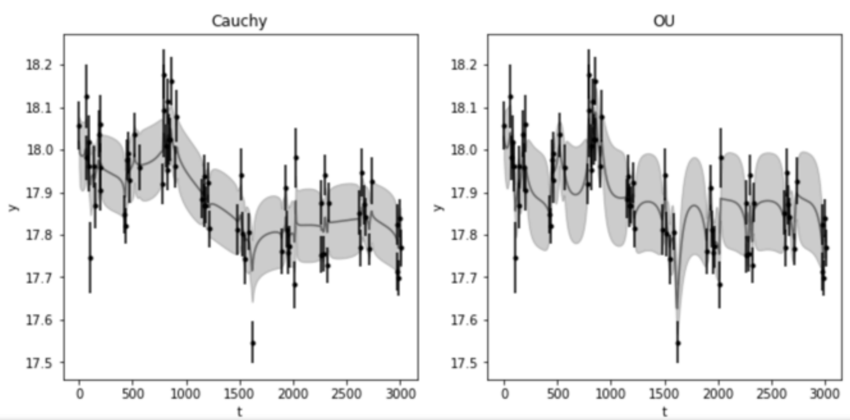


Modeling as a Gaussian process

DRW = CAR(1) = CARMA(1,0) = CARIMA(1,0,0) = CARFIMA(1,0,0)

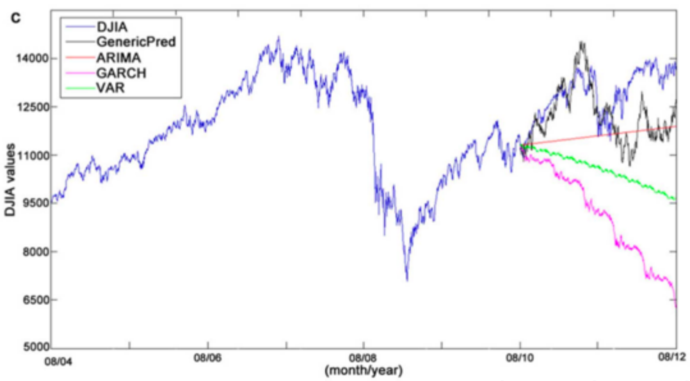
- (Zero mean) Gaussian processes are completely defined by their covariance function (kernel function)
- No closed form for (super)parent models
- Fractional Brownian motion is equivalent to CARFIMA and a Cauchy class separates characterization of the fractal dimension (roughness) and long range dependence

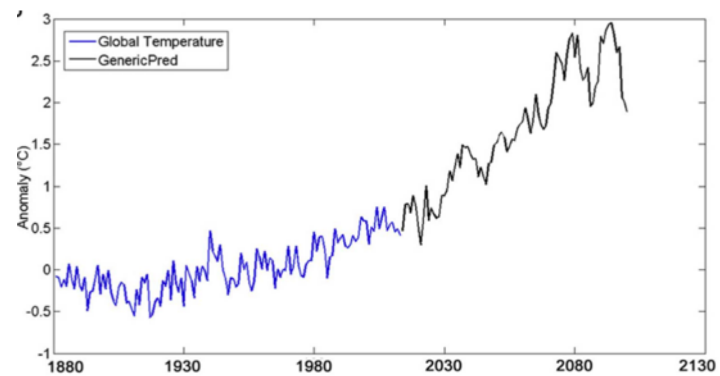
$$K(x, x') = \sigma^2 \left(1 + (|x - x'|)^{\frac{\beta}{\alpha}}\right)^{-\alpha}$$

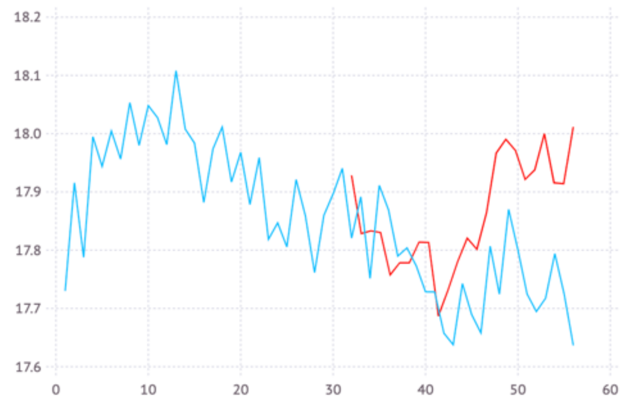


Complex nonlinear model

- Golestani & Gras (2014) proposed a method that predicts the next values in a time series such that local measures of nonlinearity, e.g., Lyapunov exponent, are as smooth as possible

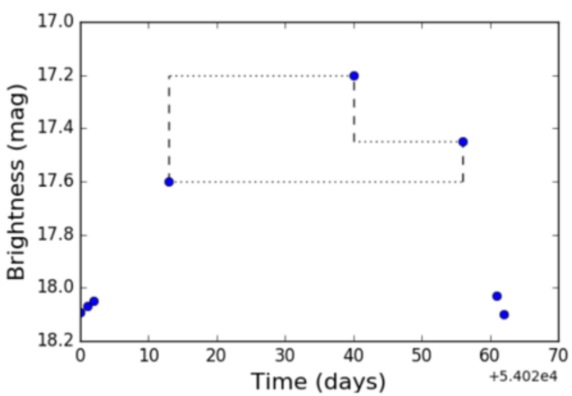


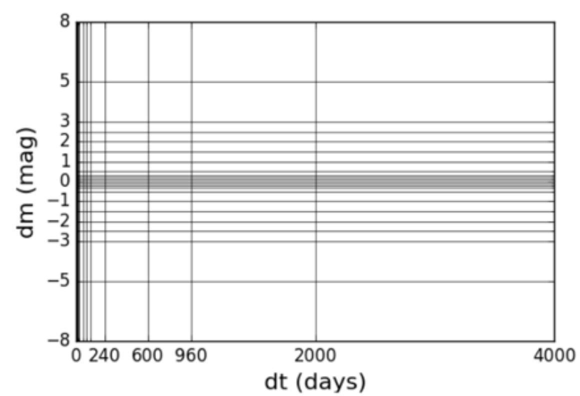


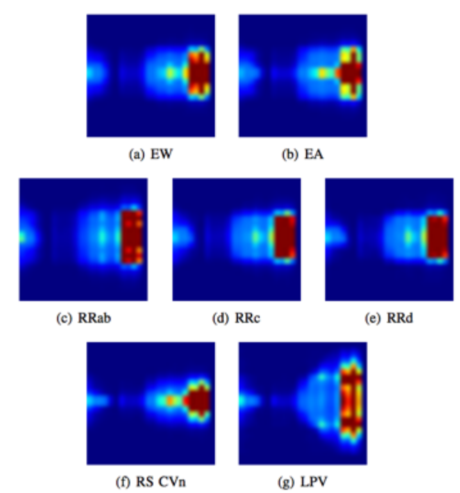


Going deep

- Trendy, good for funding proposals (preaching to the choir)
- Convolutional neural networks (CNNs) are good for images
- Need to convert time series to image:
  - Wang & Oates (2015), Hatami (2017)
  - Mahabal et al. (2017) use dm-dt mechanism with variable stars

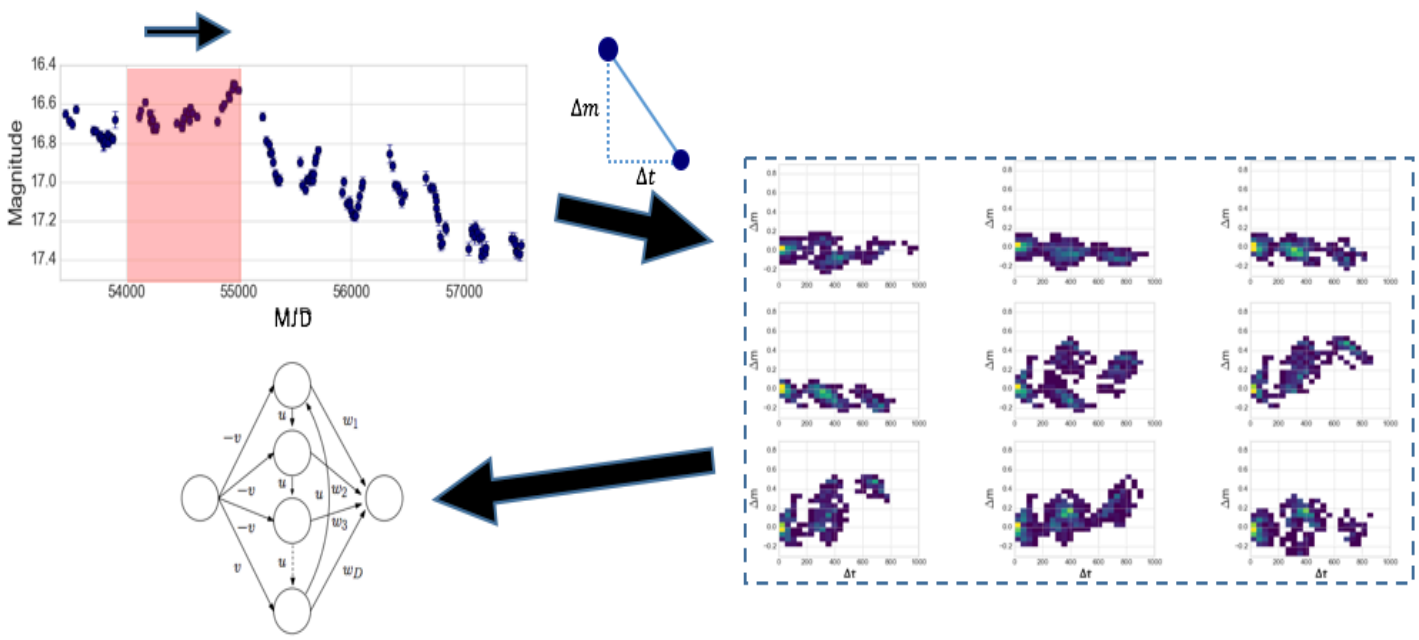






Going deeper

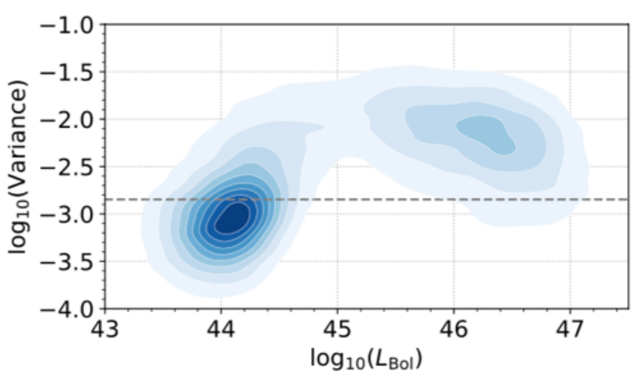
- Some neural networks architectures have “memory” - connections between links forming directed cycles, e.g., echo state, LSTMs, GRUs, etc. good for time series



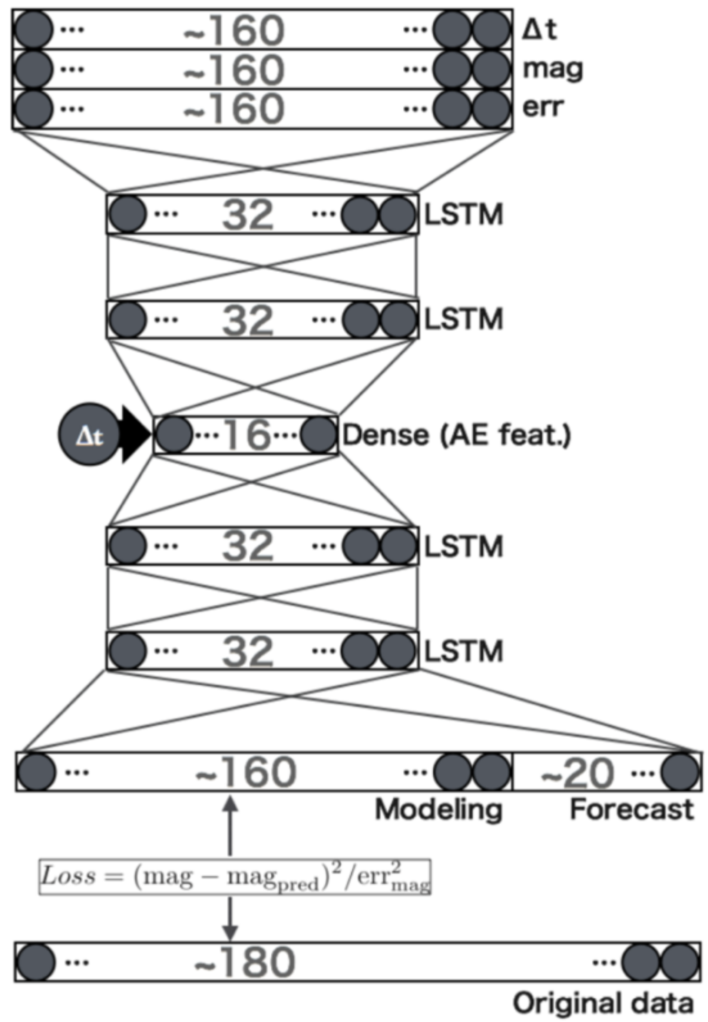
Deep modelling of time series

- Autoencoder model with RNN autoencoder:

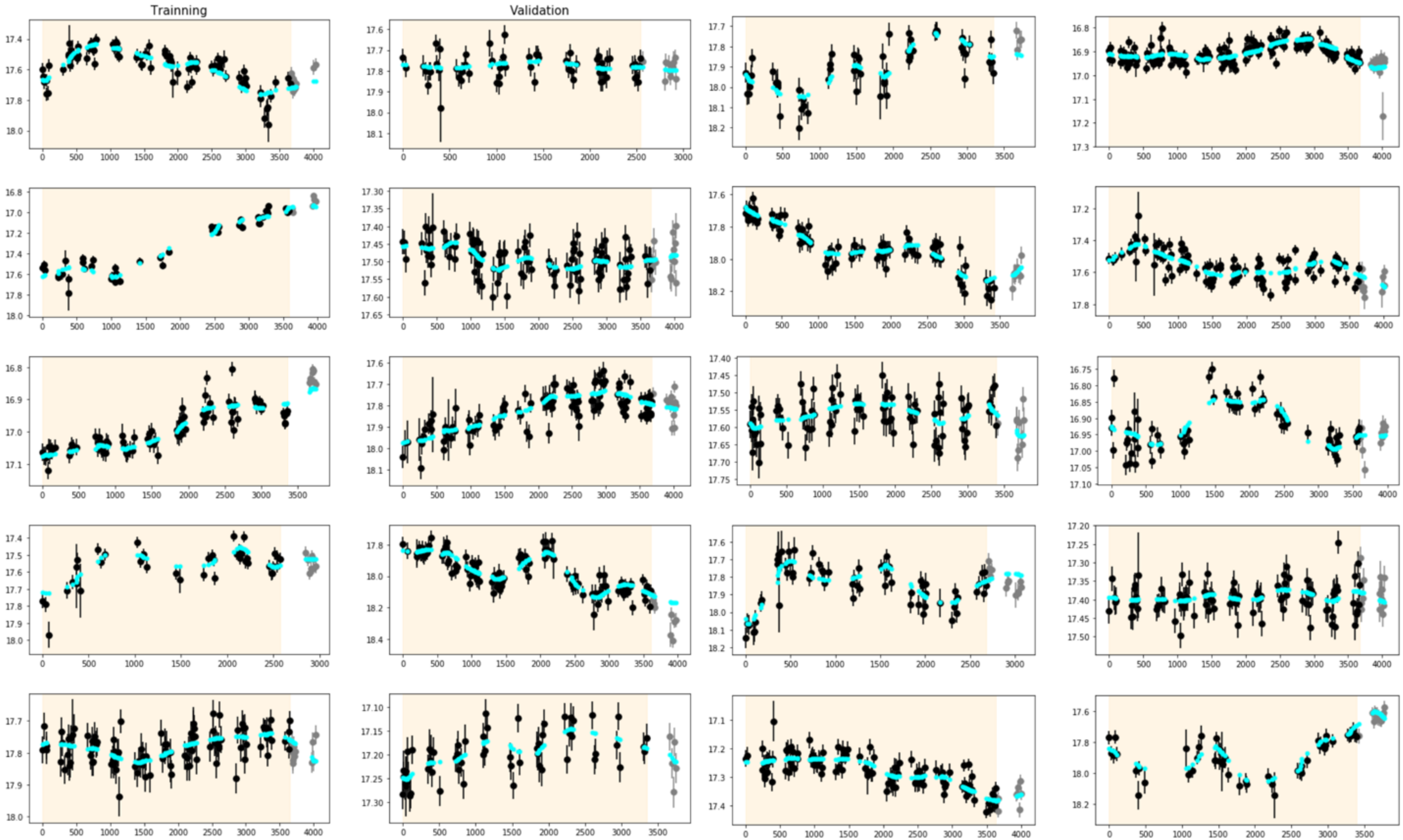
  - Consider: $(y_0, \Delta t_0) \bigoplus (\Delta t_0) \rightarrow y_1$
  - Consider: $(y_0, \Delta t_1, \dot{y}_0, \ddot{y}_0) \rightarrow y_1$



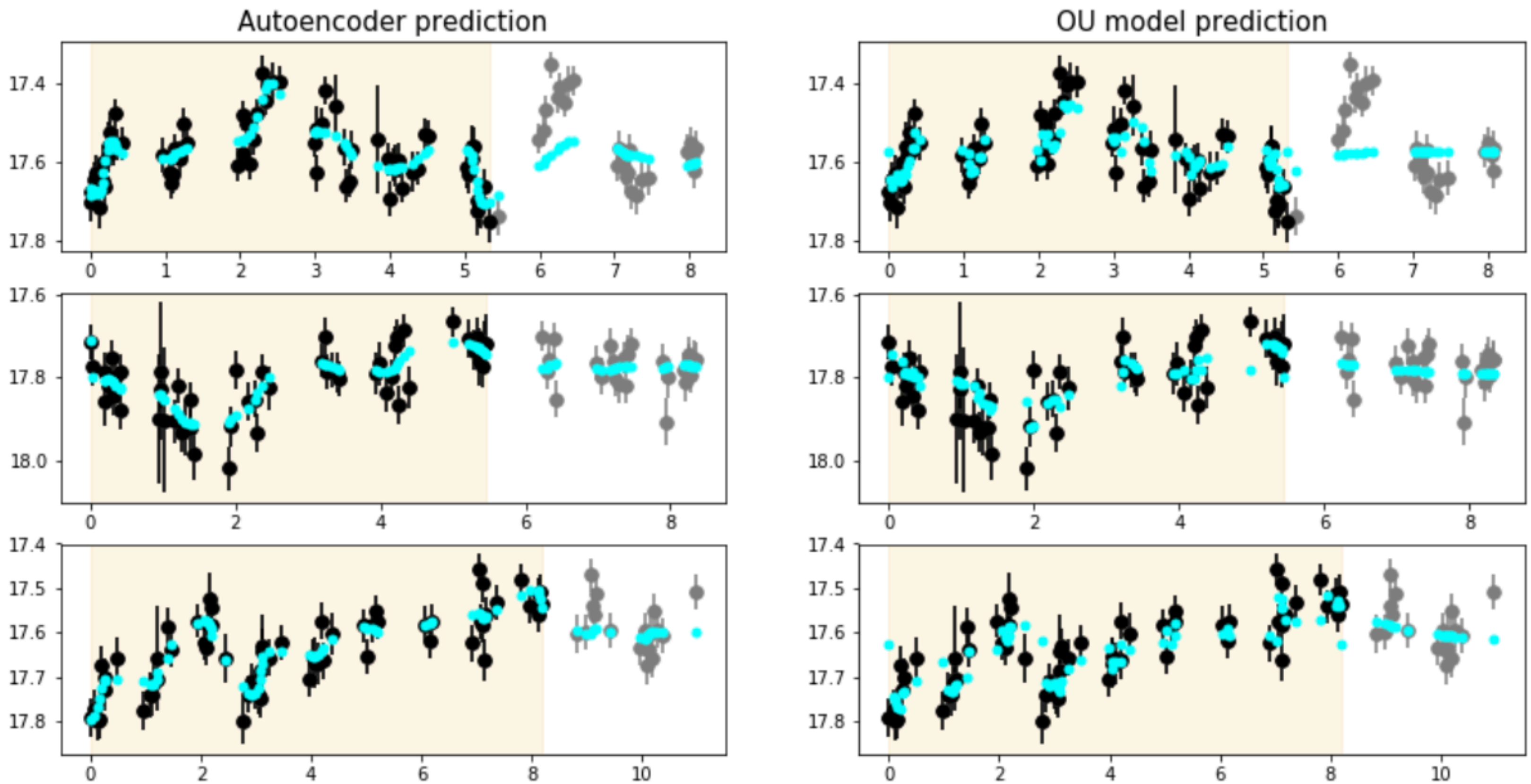
- 12,000 quasars with $\Delta t = 500$ days

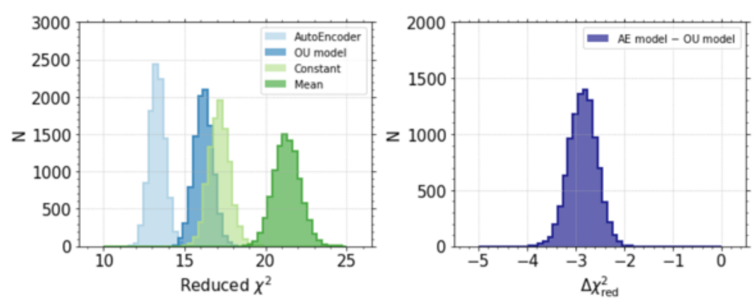


QSOs with RNNs



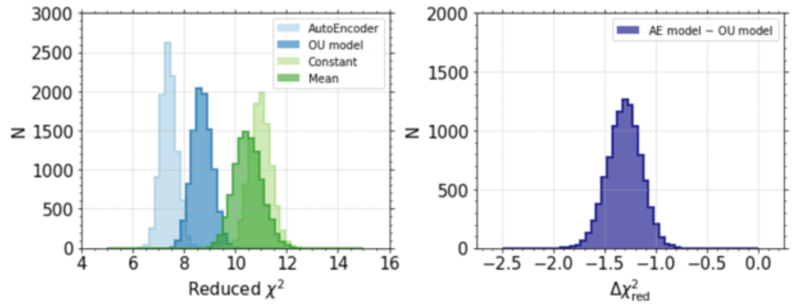
Comparing with DRW



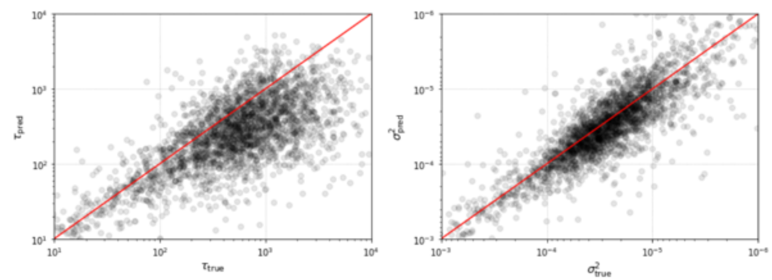


Better than DRW

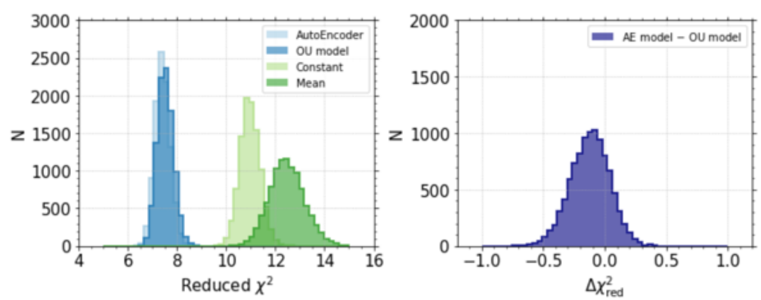
- Generate DRW light curves and fit with RNN and OU models:



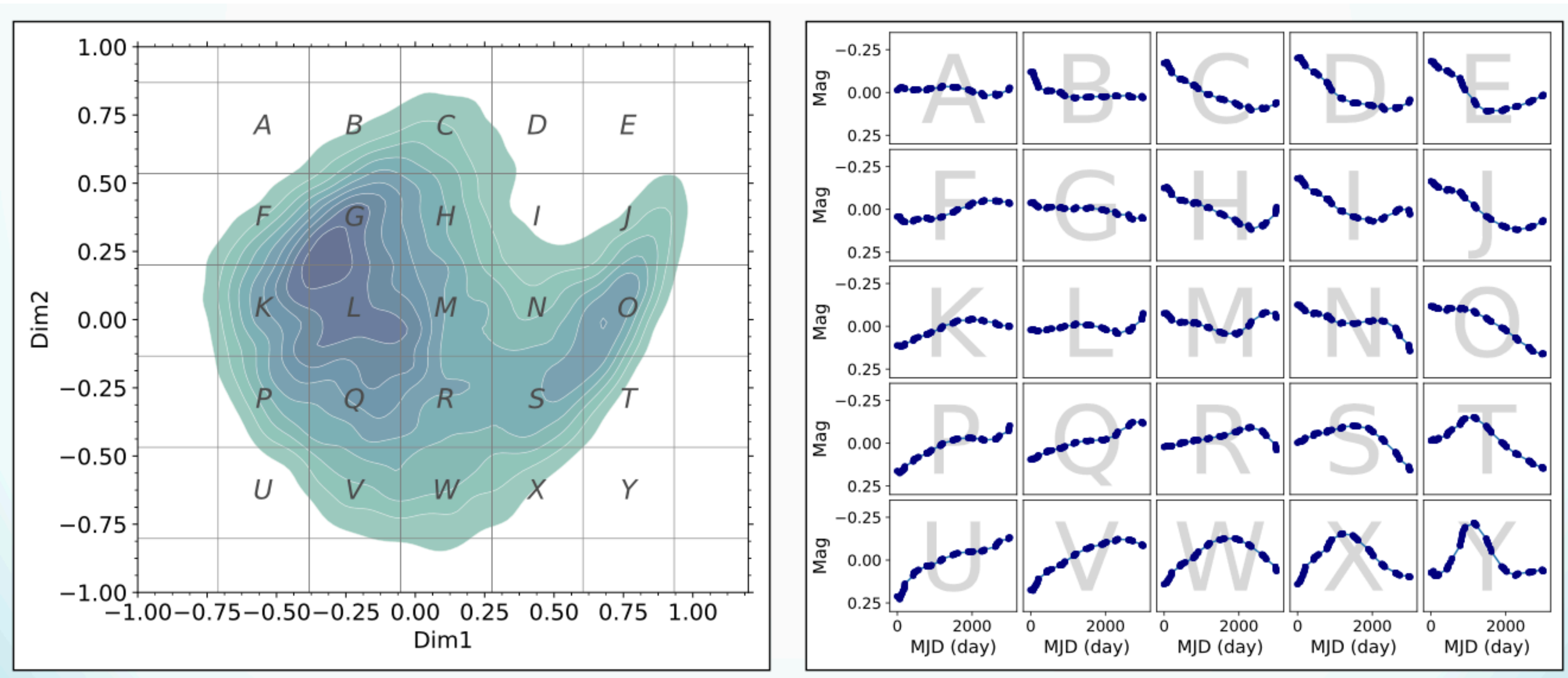
- Fitting OU models underestimates parameter values:



- True model agrees with RNN:

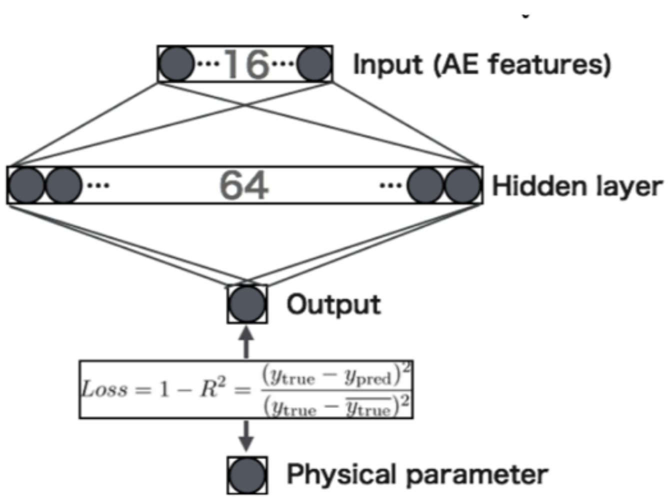


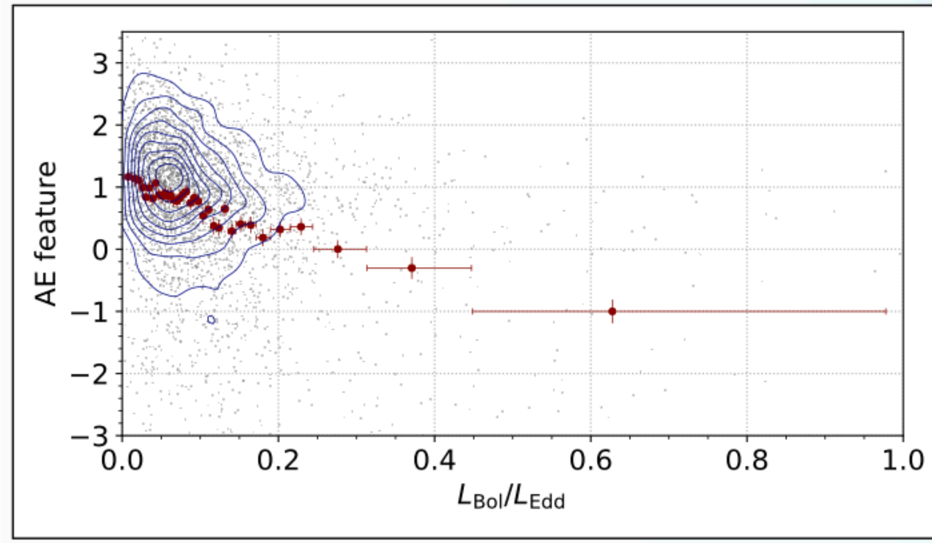
Deep time series features

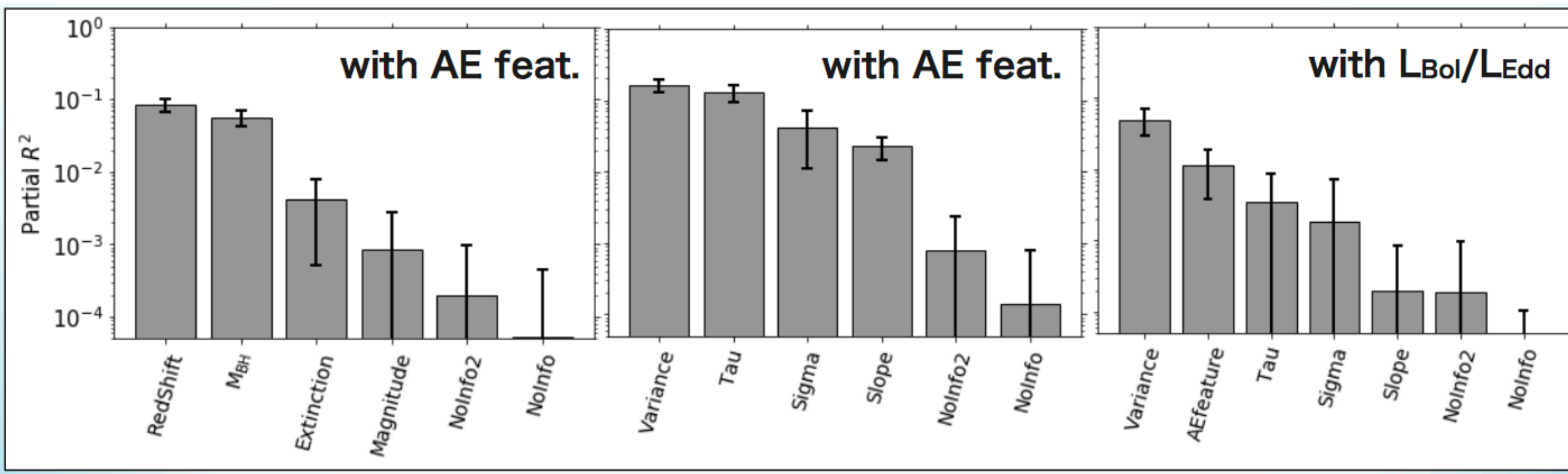


(Tachibana et al. 2019)

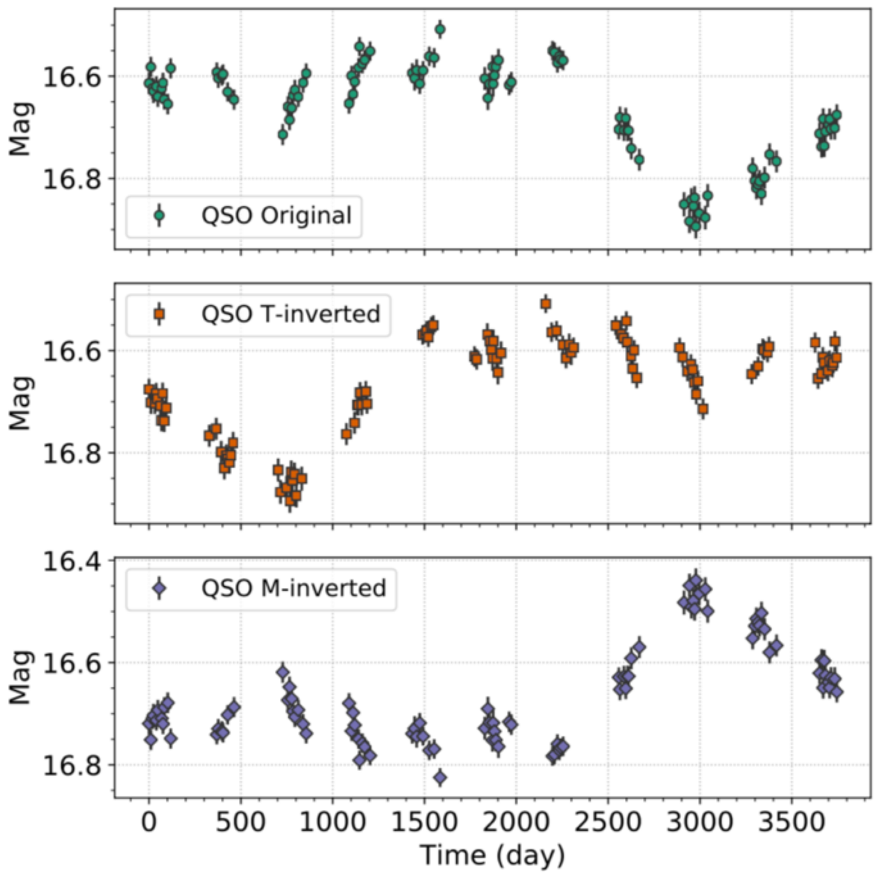
Physical correlations

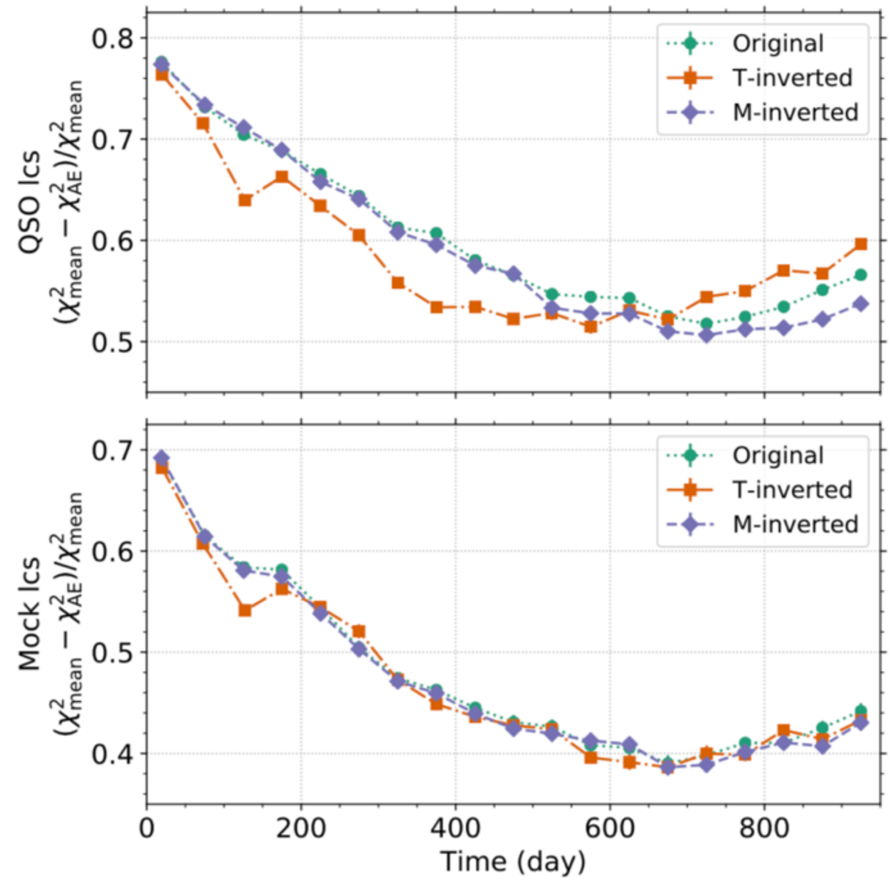






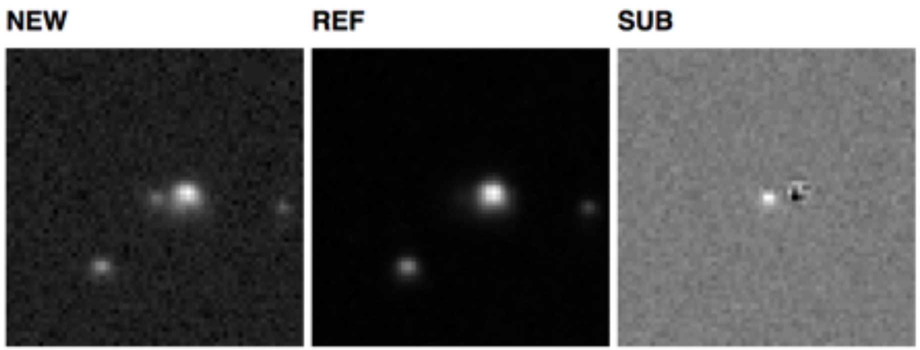
Evidence for asymmetry

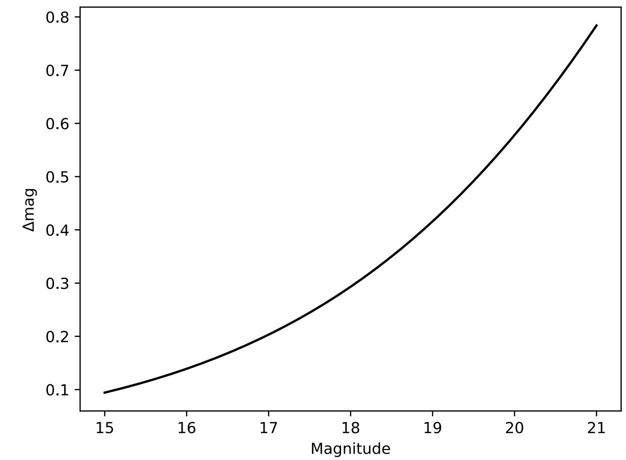


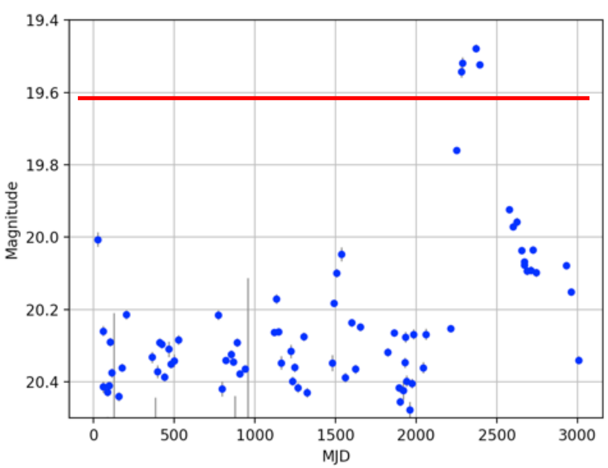


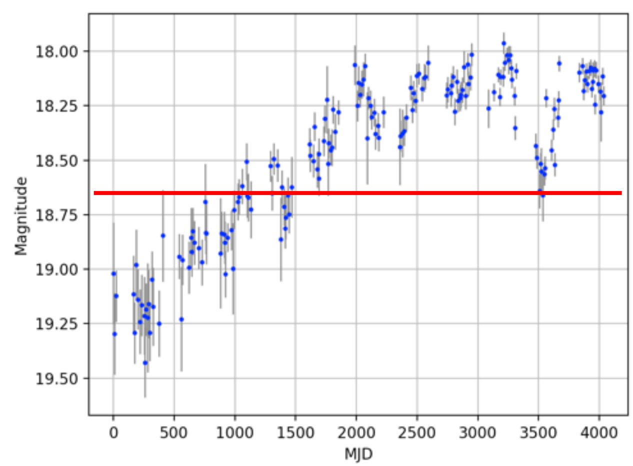
- Magnitude of asymmetry decreases as luminosity or black hole mass increases
- Consistent with self-organized disk instability model

Slow to detect as transient phenomena

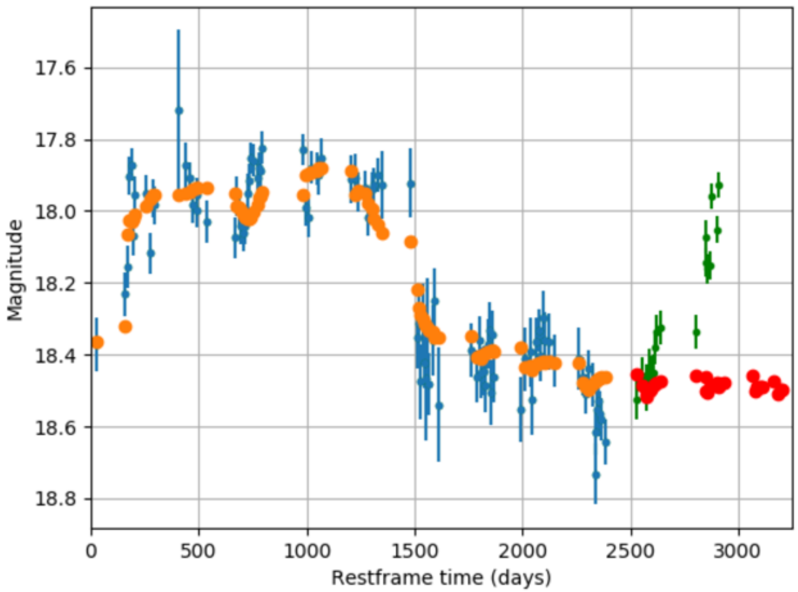


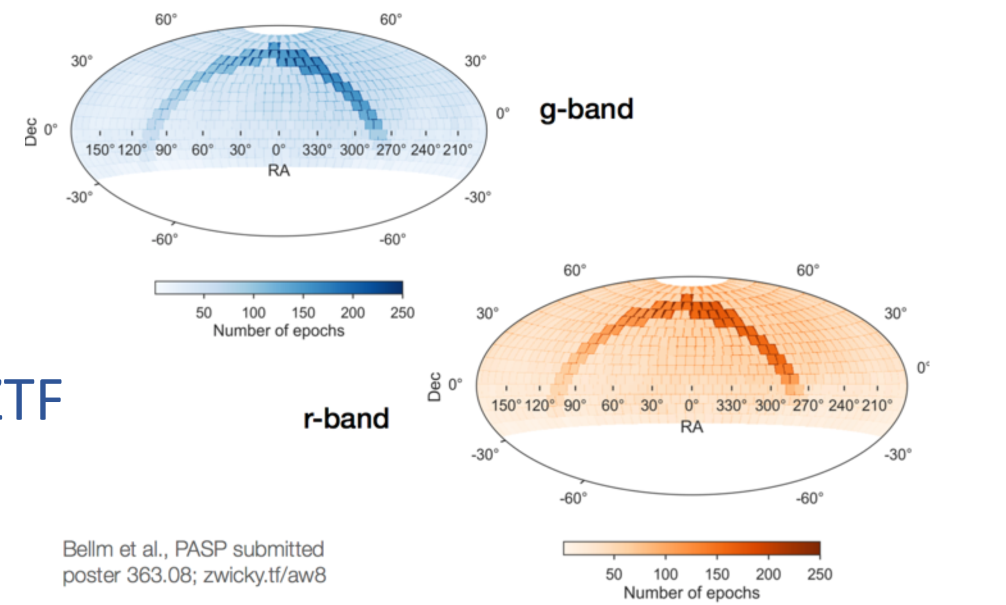






Can we forecast behavior?





- Ongoing program of looking for aberrant behavior with ZTF relative to CRTS

Bellm et al., PASP submitted
poster 363.08; zwicky.tf/aw8

Summary

- Aperiodic sources form the majority of the astrophysical population but remain the lesser studied
- Aperiodicity can be quantified
- Quasars can be modeled via Gaussian processes and smoothed local nonlinearity but RNN autoencoders provide a better model compared to DRW
- There are features which correlate with physical parameters
- The arrow of time is detectable
- Forecasting seems tractable
- Investigating other architectures:
  - Transformer, LSTM + GAN
  - Neural differential equations to handle irregular sampling
  - Asymmetry aware networks